THE VALUE OF AIR COMBAT SIMULATION

STRONG OPINIONS BUT LITTLE EVIDENCE

Wayne L. Waag

Aircrew Training Research Division
Armstrong Laboratory
Williams Air Force Base, Arizona, USA

Abstract

This paper reviews the available evidence that pertains to the value of air combat simulation. The review is limited to tactical aviation and focuses on the two principal domains of air-to-air combat and air-to-surface weapons delivery. Types of data are placed into three categories: utility evaluations that rely solely on pilot opinion data; in-simulator performance evaluations that demonstrate performance improvement as a function of training; and finally, transfer of training studies that demonstrate improved performance in the airborne environment as a function of simulation training. The primary conclusion to be drawn from the review is that although there is much opinion data that combat simulation training is valuable, transfer of training data is very limited. In fact, the only area where there appears to be solid evidence of transfer is for manual weapons delivery. There is a small amount of evidence for within visual range air combat maneuvering and none for other areas such as multi-bogey air combat and low altitude flight. Some of the difficulties inherent in the conduct of transfer of training studies within an operational military training environment are discussed. The paper then poses the question of whether there is a need to conduct to further transfer of training evaluations and concludes that such investigations may be warranted only under certain circumstances.

Introduction

In keeping with the theme of this conference, the purpose of this paper is to examine the evidence that supports the use of simulation for training tactical air combat operations. While the military has widely accepted the use of simulation for training certain types of missions and operations that are similar to commercial aviation, its value for training tactical "warfighting skills" is a highly divisive issue with vocal advocates on each side. There are some that argue that the aircraft is the _only_ environment wherein combat skills can be learned and practiced. Others argue that, even today, simulation can play an effective role in combat preparation and that its role in the future will be dramatically expanded as we are forced to accept the "peace dividend". In this paper, an attempt will be made to review the evidence that bears on this question.

The assumption that is made throughout this paper is that the acid test for determining the training effectiveness of any type of simulation,

Report Documentation Page						Form Approved OMB No. 0704-0188
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.						
1. REPORT DATE **JUN 1991**		2. REPORT TYPE **Proceedings**				3. DATES COVERED **00-01-1990 to 00-07-1991**
4. TITLE AND SUBTITLE **The Value of Air Combat Simulation Strong Opinions but Little Evidence**						5a. CONTRACT NUMBER
						5b. GRANT NUMBER
						5c. PROGRAM ELEMENT NUMBER **62202F**
6. AUTHOR(S) **Wayne Waag**						5d. PROJECT NUMBER **1123**
						5e. TASK NUMBER
						5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Armstrong Laboratory,Operations Training Research Division,6030 South Kent Street,Mesa,AZ,85212-6061**						8. PERFORMING ORGANIZATION REPORT NUMBER **AL/HR**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061**						10. SPONSOR/MONITOR'S ACRONYM(S) **AFRL; AFRL/RHA**
						11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AFRL-RH-AZ-PR-1991-0001**
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**						
13. SUPPLEMENTARY NOTES **Presented at the 1991 Royal Aeronautical Society Conference on Training Transfer - Can We Trust Flight Simulation, London UK, 1991**						
14. ABSTRACT **This paper reviews the available evidence that pertains to the value of air combat simulation. The review is limited to tactical aviation and focuses on the two principal domains of air-to-air combat and air-to surface weapons delivery. Types of data are placed into three categories; utility evaluations that rely solely on pilot opinion data; in-simulator performance evaluations that demonstrate performance improvement as a function of training; and finally, transfer of training studies that demonstrate improved performance in the airborne environment as a function of simulation training. The primary conclusion to be drawn from the training is valuable, transfer of training data is very limited. In fact, the only area where there appears to be solid evidence of transfer is for manual weapons delivery. There is a small amount of evidence for within visual range air combat maneuvering and none for other areas such as multi-bogey air combat and low altitude flight. Some of the difficulties inherent in the conduct of transfer of training studies within an operational military training environment are discussed. The paper then poses the quesstion of whether there is a need to conduct further transfer of training evaluations and concludes that such investigations may be warranted only under certain circumstances.**						
15. SUBJECT TERMS **Air combat simulation; Value; Tactical aviaition; Air-to-surface weapons delivery; Tactical training; Training effectiveness; Air combat maneuvering; Air-to-air combat training; Beyond visual range tactics; Transfer of training; Weapons delivery; Mission scenario training;**						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Public Release**	**14**		

including air combat, is transfer of training. It is assumed that the value of simulation training is measured by the extent to which, it prepares aircrews for operation in the air. While most would agree that transfer of training is a reasonable criterion for assessing the value of simulation, the difficulty usually lies in the procedures that are used to make those assessments. Caro (1977) discusses ten different approaches for estimating training effectiveness. For the purpose of this paper, I will discuss three major categories which embody these various techniques. These include utility evaluations, in-simulator learning, and transfer of training.

Utility evaluations generally make use of expert opinion as a means of estimating transfer. By far, such evaluations, are the easiest to conduct and consequently the most prevalent. What could be easier than to ask "experts" to make judgments regarding the value of training? As might be expected, however, such opinions are often suspect and of questionable validity in a scientific sense. On the other hand, such positive judgments are considered a necessary condition, although by no means sufficient for establishing transfer of training. Without pilot acceptance, the value of simulation training will likely be negligible. The second category, in-simulator learning, requires the demonstration of improved performance as a result of practice within the simulation. Simply stated, if one doesn't learn anything in the simulation environment, then transfer to the aircraft appears rather unlikely. Again, such demonstrations of improved performance are viewed as necessary, but again not sufficient for establishing training effectiveness. The final category, transfer of training, requires that improved performance be demonstrated <u>in the air</u>. If it can be shown that improved performance in the air occurs as a result of simulation training, then training effectiveness has indeed been demonstrated. In other words, such evidence is the only <u>sufficient</u> condition for establishing the effectiveness of simulation training. On the other hand as might be expected, it is also the most difficult and time-consuming evaluation to conduct.

To facilitate the review, tactical "warfighting skills" are categorized into the two general domains of air-to-air and air-to-ground combat operations. For air-to-air, it is further divided into within visual range air combat maneuvering (ACM) which emphasizes the visual aspect of air combat and the beyond visual range (BVR) multi-player, multi-bogey environment. For air-to-ground operations, there are also two categories which will be reviewed. These include conventional weapons delivery and whole mission scenarios. It must be pointed out that these distinctions are somewhat arbitrary and that combat missions often involve elements from these categories. Nonetheless, such a categorization may be of value in that they involve very different simulation requirements, especially in terms of the visual system requirements.

Air-to-Air Combat Training

Air Combat Maneuvering (ACM)

Historically, ACM has been considered the essence of air-to-air combat. It involves maneuvering to a position of offensive advantage,

achieving and recognizing entry into a weapon envelope, weapon delivery, and defensive maneuvering to defeat enemy offensive tactics. The general progression of training usually involves learning the basic elements or basic fighter maneuvers, learning to "piece together" these elements into whole engagement tactics, learning to fight together as a two-ship element, and finally learning tactics for both similar and dissimilar aircraft. To date, most air combat simulation training has focused on acquiring skill within this domain.

For ACM, there would appear to be a general consensus that indeed, simulation can provide effective training. Certainly, there has been much opinion data gathered over the years that support this view. For example, the Operational Test and Evaluation (OT&E) of the US Air Force's Simulator for Air-to-Air Combat (SAAC), originally configured to simulate the F4 Phantom, asked pilots to rate the training potential of the simulator on a task by task basis. As might be expected, a large number of tasks were judged to be trainable in the simulator. O'Neal (1984) reported similar findings for an evaluation of a limited field of view visual system for the F15. However, the most convincing opinion data have come from pilots who have received formal ACM training in such devices. Since the late 70's the USAF has conducted formal training in the SAAC. For the most part, trainees have been pilots who were already "mission-ready". The TAC ACES (Tactical Air Command Air Combat Engagement Simulation) course as it was called provided a week of intensive instruction on 1v1 and 2v1 air combat tactics. The overwhelming consensus was that the training was quite valuable.

In addition to such opinion data, studies have been conducted demonstrating significant in-simulator learning. Robinson, Eubanks and Eddowes (1981) gathered data on 34 students before and after TAC ACES training by flying two 120 second engagements against a computer controlled adversary aircraft flying preprogrammed maneuvers. The training resulted in significant improvements in weapons envelope recognition, as indicated by reductions in time until first firing opportunity, reduced time to first valid shot, increased number of valid shots, and reductions in the number of missed opportunities.

Using the same data, Eubanks and Killeen (1983) conducted further analyses by applying the theory of signal detection (TSD). This approach enabled the examination of changes in pilot performance as a function of two factors, accuracy or sensitivity (d') and response criterion or bias (β). With regard to d', training resulted in significant improvement in performance for both heat and radar missiles. This indicated a definite improvement in the pilot's ability to recognize the weapons envelope and take advantage of it. For radar missiles, β was decreased as a function of training indicating a willingness to take more shots at the end of training. For heat missiles, β increased indicating that the pilots adopted a more stringent criterion for weapons employment. The advantage of this analysis was the separation of actual improvements in performance from changes in bias or willingness to employ weapons.

McGuinness, Bouwman, and Puig (1982) also demonstrated learning effects using a measure of performance called the All-Aspect Maneuvering Index (AAMI). The AAMI is an index of positional advantage based on the

spatial relationships of opposing aircraft and the capabilities of onboard weapons. This evaluation was conducted on the US Navy's Device 2E6 which provides an air combat simulation capability for the F14. Average AAMI scores for engagements flown against a computer-driven adversary indicated improvements as a function of training. Unfortunately, no statistical analyses of these differences were reported to substantiate their reliability.

In a recent study, Raspotnik et al (1991) also demonstrated significant improvements in performance as a function of simulation training using the SAAC. Since the time of the earlier investigations, the SAAC has been modified to support both F15 and F16 training. In this investigation, two groups representing different levels of flight experience were evaluated. The results indicated significant improvements in performance due to training. These findings corroborate earlier investigations, thereby confirming that learning does occur as a function of simulation training.

We turn now to the final test, transfer of training. To date, there have been only three experimental studies, all of which were conducted in the late 70s and early 80s. Payne et al, (1976) provided simulation training to a group of eight US Navy pilots transitioning into the F4. Each pilot received 6 sorties of instruction on the Northrop Large Amplitude Simulator/Wide Angle Visual System (LAS/WAVS). Training was provided for basic fighter maneuvers such as barrel roll attacks, high yo-yo's, and rolling scissors. Performance of the simulator-trained group during normal syllabus instruction was compared against a comparable group of students receiving no simulation training. Results indicated that the simulator-trained group achieved superior final position outcomes during engagements flown in the air. Likewise, they also received higher grades by their instructors. Such group differences were maintained throughout the entire tactics syllabus of instruction.

Pohlmann and Reed (1978) conducted a study in the SAAC designed to evaluate the contribution of platform motion to the initial acquisition of BFM skills. Sixteen students received 7 training sorties in the SAAC; half with platform motion, the other half without. Their performance on two aircraft "data rides" was compared against a control group (N=6) who did not receive SAAC training. An analysis of the data collected in the air revealed no improvement in performance as a result of simulator pretraining. In fact, the trend was toward better performance by the control group.

A second study conducted in the SAAC (Jenkins, 1982) investigated the effects of air-to-air simulation training on subsequent performance during a Fighter Weapons Instructor Course (FWIC) conducted at Nellis AFB, Nevada. The performance of 14 pilots receiving training in the SAAC was compared against the performance of 14 pilots receiving no simulation training. Gun camera film was analyzed to determine the number of attempted and valid missile shots and gunshots taken during the FWIC sorties . The results indicated the SAAC-trained group achieved a higher average percentage of valid missile shots and valid gunshots. They also achieved higher exchange ratios and achieved a higher class standing in the course.

Although not a controlled investigation, some mention should be made of an evaluation conducted by the RAF's 228 Operational Conversion Unit (OCU) in the early 80s. Over a one week period, four pilots converting to the F4 received 7 sorties of instruction in the BAe Warton Twin Dome Air Combat Simulator. Written performance reports obtained following the simulator exercises were compared against reports following the actual air exercises. The conclusion was that the simulator-trained group progressed noticeably faster when compared against previous courses who had not received the simulation training.

Taken as a whole, the evidence is fairly positive toward the use of simulation for training ACM. Certainly, the opinion data obtained from utility evaluations is quite positive. Moreover, the data from in-simulator learning studies suggests that performance does reliably improve as a function of training. Unfortunately, the data from the transfer studies are quite limited. Of the three studies to date, two have produced positive effects while one has shown no effect of the training. It should be pointed out that the one investigation showing no effect (Pohlmann and Reed, 1978) made use of instructor ratings for assessing performance both during simulation training and the two aircraft "data rides". It is quite likely that this finding of no effect may have been largely due to the insensitivity of such measures which also failed to show consistent learning effects during the simulation training. We shall return to a discussion of measurement problems later in this paper.

Beyond Visual Range (BVR) Tactics

Within visual range ACM training is concerned primarily with individual skills training. In fact, the literature cited in the previous section has dealt exclusively with training for the 1v1 environment. While individual skills are important, the fact is that the basic fighting element is the two-ship. Moreover, as weapons systems have become increasing sophisticated, there has occurred a change in doctrine toward reliance upon beyond visual range capabilities and the use of medium and long range missiles. The shift in emphasis toward beyond visual range tactics has led to questions of the value of simulation training for this environment.

Before reviewing the evidence to date, it may be of value to describe the salient characteristics of the BVR environment. In a nutshell, the key element is complexity. It is characterized by multiple players, both friendly and foe. Players include not only the pilots but elements of command and control such as GCI and AWACS controllers. It is also a highly intensive electronic environment with the requirement that onboard radar and avionics systems be accurately modeled. There are multiple threats, both on the ground and in the air. There are also elements of terrain and weather that must be modeled. In sum, the emphasis of such simulation training is generally placed upon the environmental and situational fidelity.

Although the concept of multiplayer air combat simulation training is not new (E.g. Hughes and Brown, 1984; Hughes, etal, 1985), it is only recently that efforts have been initiated to explore the value of such training. In 1988, the Armstrong Laboratory (then called the Air Force

Human Resources Laboratory or AFHRL) initiated a program with the Tactical Air Command to evaluate multiship air combat training using commercially-available contractor facilities (Thomas, Houck, and Bell, 1990). Forty-two mission-ready F15 pilots and 16 GCI/AWACS controllers were each trained over a four day period at the McDonnel Douglas simulation facility at St Louis, Missouri. The basic unit of training was the two ship element (lead/wingman) plus the controller against an opposing force comprised of 4 to 8 adversaries plus the adversary controller. Mission scenarios included sweep vs sweep, combat air patrol, and force protection.

At the beginning of the training period, pilots rated the "training need" for 41 tasks. Upon completion of training, pilots rated the value of both their "unit training" and the "simulation training" for the same 41 tasks. The results indicated that for certain tasks, training in the simulator was considered much better than the actual training in the air received at their unit. Representative tasks included multibogey, chaff and flares employment, all-aspect defense, ECM/ECCM employment, communications jamming, etc. In other words, they were tasks that were rarely practiced in their home units. On the other hand, such things as ACM, visual lookout, gun employment, and BFM were better trained at their home units. In other words, they were tasks that were practiced most often in their home units. Interestingly enough, the controllers rated all tasks to be trained better in the simulation environment. Open-ended opinion data were also gathered, the results being quite positive toward the training.

Houck, Thomas, and Bell (1991) conducted a follow-on evaluation using the same procedure but with a larger and more representative sample of pilots and controllers. Similar results were obtained in that there were tasks trained in the simulator that were rated as more valuable than training in the home unit. Likewise, the opinions expressed were extremely positive. As a result of these efforts, and perhaps the most convincing evidence for user acceptance, TAC has continued this program under its own sponsorship.

However, positive user opinion is but one of the conditions for establishing transfer of training. In 1990, a study was conducted again at the McDonnel Douglas simulation facility with the intent of determining in-simulator learning effects. Subjects consisted of 16 elements (32 pilots) and controllers. An attempt was made to control experience level with four of the elements representing "expert" levels with the remaining 12 representing "journeyman" levels. Each of the elements flew controlled offensive and defensive scenarios "before" and "after" 4 days of intensive simulation training. Digital data as well as videotapes of displays used for replay and debriefing purposes were archived and are currently being analyzed. Initial results indicate that significant improvements in performance did occur for certain outcome events such as kills and exchange ratios. While such results are only tentative, it seems clear that learning does occur.

To summarize the evidence to date, it is clear that user opinion is quite favorable to simulation training for the multiship BVR environment. Moreover, there is some data suggesting that sigifnicant learning does

occur. However, at present transfer of training data have not been gathered and therefore the sufficient condition has not been established.

Air-to-Surface Combat Training

Weapons Delivery

Weapons delivery or dropping bombs is generally considered the terminal event of most ground attack missions. It is also an area where the effectiveness of combat simulation training has perhaps been most clearly demonstrated. Beginning in the mid 70s, a number of studies were conducted which produced positive transfer of training results. Of particular interest was a study by Gray and Fuller (1977) which evaluated the transfer of weapon delivery training using the Advanced Simulator for Pilot Training (ASPT) located at AFHRL at Williams AFB, Arizona. At that time, the ASPT was configured to simulate the T37 aircraft, the USAF's primary jet trainer. The original intent of the study was to investigate the effects of platform motion cuing so that two groups of 8 students receiving training in the ASPT were compared against a group of 8 students who had not received simulation training. The aspect of most interest was the fact that although training was accomplished in a T37 simulation (note that a fixed gunsight was added), the actual transfer evaluation was conducted in the F5B aircraft. Student pilots receiving training in the ASPT scored significantly better on all measures of bombing accuracy as compared to the group of students who had not received the pretraining.

A subsequent study by Gray, Chun, Warner, and Eubanks (1981) produced similar findings. By that time the ASPT had modified to an A10 configuration and was being used by students transitioning into that aircraft. Seventeen students received three sorties of simulation training in conventional weapons deliveries, pop-up deliveries, and low angle strafing. Subsequent performance in the aircraft was compared against a group of 7 students not receiving simulator pre-training. Significant transfer was demonstrated in four of the five conventional deliveries, pop-ups, and the strafing event. These differences were most pronounced on the first three or four sorties in the aircraft.

Hagin, Durall, and Prophet (1979) evaluated the effectiveness of weapons delivery training using the US Navy's Device 2B35/2F90. Students received four sorties of training in the simulator in which emphasis was placed upon setting up the correct pattern and releasing weapons at correct parameters. The performance of this group was compared against students not receiving the simulation training. Results indicated significantly fewer pattern errors, although the groups did not differ significantly in either simple bomb miss distance or the horizontal or vertical components.

Based on the available evidence, it seems pretty clear that positive transfer can be expected for conventional weapons delivery training. It is of some interest that the studies to date were all accomplished with aircraft requiring manual weapons delivery. The extent to which these results generalize to most newer weapons systems which provide computer-aiding is unknown.

Mission Scenario Training

While weapons delivery is an essential part of the surface attack mission, it represents only a small portion of the total training requirement to correctly execute the entire scenario. Most interdiction or close air support missions entail navigation to the target area, usually at low altitude, ingress into the target area, attack/re-attack, egress from the target area, and navigation back to the home base. Throughout the mission, perhaps the greatest concern is with threats, both ground and airborne. Learning defensive tactics necessary to defeat those threats and survive has been considered a critical training problem.

To date, the only controlled studies in this arena have been conducted by AFHRL. With the conversion of the ASPT to an A10 configuration in the late 70s, interest arose in the potential value of simulation for teaching defensive tactics within a high-threat environment. This seemed to be a reasonable question following the demonstration of effective training in the A10 for weapons delivery. In the first investigation, Kellogg, Prather, and Castore (1980) reported significant in-simulator practice effects within a high threat environment. Upon completion of training, mission success was improved both in terms of targets destroyed, as well as survivability against ground threats.

A series of transfer studies were then initiated to determine whether simulation training in a high threat environment would have some effect upon subsequent performance in operational exercises such as Red Flag. In the first study, Hughes, Brooks, Graham, Sheen, and Dickens (1982) provided simulation pre-training for a group of 11 mission-ready A10 pilots who were scheduled for Red Flag 82-2. Each pilot received two hours of simulation practice in both battlefield interdiction and close air support missions. Pilots were allowed to simply "fly" the missions without any attempt to structure the exercises. Performance at Red Flag was compared against a group of 14 A10 pilots who had not received the simulation pre-training. The results indicated a significant increase in survivability for the simulator-trained group in which the RWR/ECM configuration of the aircraft flown at Red Flag was the same as the simulataor configuration. However, for the group in which the configuration was different, there occurred a decrease in survivability.

The importance of these findings lies in two areas. First, these results clearly demonstrated a positive transfer effect as a result of the simulation training when the avionics configurations were the same. Equally important was the demonstration of negative transfer when the configurations were different. In addition to these findings, opinion data were gathered regarding the training. In general, opinions were positive except it was universally agreed that the visual system was in need of improvement, especially in the areas of resolution and need for scene detail near the ground.

In a follow-on study, Wiekhorst and Killion (1986) provided simulation pre-training in the same hostile environment to 13 A10 pilots prior to their participation in the Green Flag 84 exercise. Their performance was compared against 38 A10 pilots who had not received pre-training within the ASPT hostile environment simulation. The results were consistent with

previous findings demonstrating improved performance during the flag exercise in terms of both survivability and more effective use of self-protection countermeasures.

Other than these two transfer studies conducted for the A10, there is little hard evidence supporting the value of combat simulation for training surface attack mission scenarios. However, in 1989 the Armstrong Laboratory conducted a feasibility demonstration of two-ship training for the F16 at the General Dynamics simulation facility located at Ft Worth, Texas. The demonstration was conducted over a two week period in which 12 mission-ready pilots (6 elements) flew a variety of interdiction and close air support ground attack missions as well as two-ship defensive and offensive air-to-air missions. The consensus of those participating in the demonstration was that there is significant training potential for simulation training for both the ground attack and air-to-air environments.

Discussion

We turn now to the original question posed at the beginning of this paper--how convincing is the evidence that supports the use of simulation for training tactical air combat operations? It should be apparent to the reader that there exist very limited hard data, i.e. transfer of training data. Most of the studies were conducted in the late 70s and early 80s, and interestingly enough most were conducted by Armstrong Laboratory. While there exist a fair amount of opinion data that suggests there is training potential in using simulation, the data that would be considered sufficient to substantiate that potential is extremely limited. For example, in the whole domain of air-to-air including both ACM and multiship BVR tactics, only three transfer studies have been conducted. Of these only two produced positive results, and a careful reading of the actual reports suggest that the size of the effects, even though significant, were fairly small. For the surface attack training, the results, at least for conventional weapons delivery, are somewhat more encouraging. It seems fairly well established that simulation can be effectiveness for training these events. However, it should be pointed out that manual deliveries are no longer trained for most current generation weapons systems. The two studies demonstrating transfer to the Flag exercises are perhaps the most encouraging, but again the effects, although significant were fairly small.

So what's the answer? Is simulation training for tactical combat operations effective in terms of transfer to the aircraft? The answer is "probably yes", but for reasons only partially related to the results of these transfer tests. In attempting to articulate this position it may be instructive to discuss the transfer of training methodology and some of the problems of implementation within the military training environment. I have discussed many of these problems in a report prepared some years ago (Waag, 1981) and will only briefly mention them here. Lack of experimental control, insufficient sample sizes, insufficient training time in the simulator, insufficient time for evaluating transfer in the air, insensitive measures, etc. are all problems that plague the conduct of any transfer of training evaluation. In fact, one can argue that it is virtually impossible to conduct a well controlled transfer test within an

operational military environment.

The impact of this inability to adequately control such evaluations is perhaps greatest upon the interpretation of findings, in particular findings of no transfer effects or fairly small transfer effects, which would certainly characterize the data within the domain of combat simulation training. In such instances of no effects or very small effects, there are two possibilities. Indeed, there may be no effects or very small effects. Or, the effects may be much larger, but because of the "problems" inherent in conducting such evaluations, they are masked. In such instances, we generally attempt to "explain away" the lack of positive effects and attribute it to these "problems", especially if there are other data such as expert opinion that suggest the training to be beneficial.

A good case in point is the study by Pohlmann and Reed (1978) that failed to show any positive transfer effects. Do we accept these findings? Probably not, since we "know" the training to be beneficial because of all the positive end-of-course critiques indicating that such training in the SAAC was some of the best air-to-air training pilots had ever received. Moreover, the study had a serious limitation in that the measures of performance were instructor ratings which are notoriously insensitive. For example, the study by Gray and Fuller (1977) which demonstrated significant transfer of training in terms of bombing accuracy, also used instructor ratings of performance in the air. Interestingly enough, the rating data showed no effects of simulator pre-training despite the fact of large differences in objective measures of weapons delivery. So it seems at least plausible that the failure to show any effect in the Pohlmann and Reed (1978) study was due largely to the measures that were used. For this reason and the fact that we "know" the training to be valuable, we can make the case to simply "dismiss" these findings.

At this point, the reader should begin to see a paradox emerging. On the one hand, we have made the argument that the transfer of training evaluation is the only <u>sufficient</u> test for establishing training effectiveness. On the other hand, we have also shown that we tend to dismiss those studies failing to demonstrate positive transfer when we have other data, which is usually expert opinion, suggesting the training to be effective. In such instances we attribute the lack of positive transfer effects to one or more of those "problems" which always exist in the conduct of such evaluations within an operational military training environment. If such is the case, the question becomes, "why conduct the transfer evaluation?".

To illustrate the dilemma, consider the recent efforts by the Armstrong Laboratory in training multiship BVR tactics using the McDonnel Douglas simulation facility. As previously discussed, user opinion has been quite positive. The training program has now transitioned to the Tactical Air Command. Data gathered in controlled evaluations indicate that significant learning does occur in terms of improved kills, exchange ratios, etc. In other words, performance is improved and aircrews express positive opinions toward the training. However, these are only necessary but not sufficient conditions for establishing training effectiveness. To

date, we have not demonstrated that what is learned during such training actually transfers to the air. The question becomes, "should we?".

I'm sure the reader will understand the difficulties involved in such an evaluation. First of all, since we are talking a 2v4 environment, it should be clear that such an evaluation would be expensive. Second, even if such an evaluation were conducted on an instrumented range, the amount of data available is limited primarily to aircraft state and position information. And third, there is the problem of assessing performance of tasks that are simply not done in the air. For example, effects of communication jamming can be practiced in the simulator but not in the air due to safety constraints. Despite these difficulties, suppose that we were able to gather the necessary support to conduct at least some semblance of a transfer of training evaluation, probably similar to the Flag exercise studies conducted in the early 80s. At best, we can probably expect very limited effects. If on the other hand we show no effect (which in my view would be quite likely), do we conclude the simulation training to be of no value? Again, "probably not". We would most likely attribute the failure to demonstrate positive effects to all the limitations that will inevitably characterize the evaluation.

The question remains, "do we conduct the transfer test?". In my view, the answer is "no", at least for the multiship BVR environment. In part, this answer is based upon the argument just presented. However, I think there are other reasons to suggest that such evaluations may be unnecessary, at least within this environment. If one looks at the whole of the simulation training effectiveness literature, the overwhelming conclusion is that transfer of training does occur. In particular tasks learned that are of a procedural nature have been found to transfer quite well. Since modern fighters require less psychomotor skill and more procedural/systems management skills, it seems likely that transfer should be high. Moreover, theory suggests that the more closely that the simulation environment resembles the actual environment, the greater the transfer. Again, since today's simulators can easily model avionics and weapons functions quite accuracy, one can expect a high degree of transfer. In sum, there seems little reason <u>not</u> to expect positive transfer in the multiship BVR environment.

On the other hand, there may be other environments where such arguments would not apply. For example, low altitude flight is an area that has been identified as a candidate for visual simulation training. Unlike BVR tactics, there is a very strong perceptual-motor component in low altitude flight. The capability of simulation to support low altitude flight training is certainly questionable. The only transfer study to date which focused on low level navigation (Pierce, 1983) failed to show any effect. This result coupled with the generally "negative" opinion regarding the capabilities of visual simulation at low altitude would certainly suggest that a transfer evaluation should be undertaken at some point in time. The point to be made from this discussion is that although transfer of training is indeed the acid test for training effectiveness, there may be no need to conduct such an evaluation for certain environments that are similar to those, in terms of relevant behavioral components, for which transfer has previously been shown.

In my review of the training effectiveness literature 10 years ago (Waag, 1981), I concluded that one of the major issues was generalization. In other words, to what extent does one set of findings generalize to other situations? The state of the literature then and certainly today, is that applied researchers have always conducted transfer of training evaluations within a specific context since they were usually attempting to answer some "real-world" question. The problem I saw then, and that still exists today, is that we lack a basic understanding of the transfer process at a more fundamental level of behavior. If indeed we had a better understanding of skill acquisition and transfer at such a level, there would be no need to conduct these time-consuming, expensive, and often inconclusive transfer investigations. It should be possible to analyze training requirements into their fundamental skill components and based upon our understanding of the transfer process at the component level, generate predictions of real-world simulator transfer. Unfortunately, I am unaware of any significant research programs that are addressing these fundamental issues. Until such work is done, I suspect that we will continue in our current mode of asking questions of transfer for each new simulator that is fielded for each new application.

References

Caro, P. (1977). Some factors influencing Air Force simulator training effectiveness. (HumRRO-TR-77-2). Alexandria VA: Human Resources Research Organization.

Eubanks, J.L. and Killeen, P.R. (1983). An application of signal detection theory to air combat training. Human Factors, 25, 449-456.

Gray, T.H. and Fuller, R.R. (1977). Effects of simulator training and platform motion on air-to-surface weapons delivery training. (AFHRL-TR-77-29). Williams AFB AZ: Air Force Human Resources Laboratory.

Gray, T.H., Chun, E.K., Warner, H.D. and Eubanks, J.L. (1981). Advanced flight simulator: Utilization in A-10 conversion and air-to-surface attack training. (AFHRL-TR-80-20). Williams AFB AZ: Air Force Human Resources Laboratory.

Hagin, W. V., Dural, E. and Prophet, W.W. (1979). Transfer of training effectiveness evaluation: US Navy device 2B35. (Seville Research Corporation TR 79-06). Pensacola FL: Chief of Naval Education and Training.

Houck, M.R., Thomas, G.S. and Bell, H.H. (1991). Training evaluation of the F-15 advanced air combat simulation. (AL-TP-1991-0047). Williams AFB AZ: Armstrong Laboratory.

Hughes, R., Brooks, R.B., Graham, D., Sheen, R. and Dickens, T. (1982). Tactical ground attack: On the transfer of training from flight simulator to operational red flag exercise. Proceedings of the Human Factors Society, 26 Annual Meeting.

Hughes, R.G., and Brown, L. (1984). Trends shaping advanced aircrew training capabilities through the 1990's. (AFHRL-TP-84-52). Williams AFB AZ: Air Force Human Resources Laboratory.

Hughes, R.G., Polis, D., Fay, R.H., Hines, J., and Altman, H. (1985). Tactical training complex: Prototype facility for integration of advanced simulation and range system concepts for tactical aircrew training. (AFHRL-TR-85-3). Williams AFB AZ: Air Force Human Resources Laboratory.

Jenkins, D.H. (1982). Simulator training effectiveness evaluation (TAC Project 79Y-001F). Nellis AFB NV: USAF Tactical Fighter Weapons Center.

Kellogg, R., Prather, E. and Castore, C. (1980). Simulated A-10 Combat Environment. Proceedings of the Human Factors Society, 24th Annual Meeting, pp 573-577, Los Angeles, CA.

McGuiness, J., Bouwman, J.H. and Puig, J.A. (1982). Effectiveness evaluation for air combat training. In Proceedings of the 4th Interservice/Industry Training Equipment Conference: Volume I. Washington DC: National Security Industrial Association.

O'Neal, M.E. (1984). F-15 limited field of view visual system training effectiveness evaluation. (TAC Project 83G-066T). Eglin AFB FL: USAF Tactical Air Warfare Center.

Payne, T.A., Hirsch, D.L., Semple, C.A., Farmer, J.R., Spring, W.G., Sanders, M.S., Wimer, C.A., Carter, V.E., and Hu, A. (1976). Experiments to evaluate advanced flight simulation in air combat pilot training. Vol 1. Transfer of learning experiment. Hawthorne CA: Northrop Corporation.

Pierce, B.J. (1983). Low altitude simulator training: A-10 aircraft. (AFHRL-TR-82-43). Williams AFB AZ: Air Force Human Resources Laboratory.

Pohlmann, L.D. and Reed, J.C. (1978). Air-to-air combat skills: contribution of platform to initial training (AFHRL-TR-78-53). Williams AFB AZ: Air Force Human Resources Laboratory.

Raspotnik, W.B., Leeds, J., Brecke, F., Miller, D., Hayes, E., Thomas, G., Dixon, K., and Waag, W.L. (1991). Validation and refinement of techniques for air combat assessment. In preparation. Williams AFB AZ: Armstrong Laboratory.

Robinson, J.C., Eubanks, J.L. and Eddowes, E.E. (1981). Evaluation of pilot air combat maneuvering performance changes during TAC ACES training. Nellis AFB NV: USAF Tactical Fighter Weapons Center.

Thomas, G.S., Houck, M.R. and Bell, H.H. (1990). Training evaluation of air combat simulation. (AFHRL-TR-90-30). Williams AFB AZ: Air Force Human Resources Laboratory.

Waag, W.L. (1981) Training effectiveness of visual and motion simulation (AFHRL-TR-79-72). Williams AFB AZ: Air Force Human Resources Laboratory.

Wiekhorst, L.A. and Killion, T.H. (1986). Transfer of electronic combat

skills from a flight simulator to the aircraft. (AFHRL-TR-86-45). Williams AFB AZ: Air Force Human Resources Laboratory.